CDF Note 11025

Measurement of Single Top s-channel Cross Section in Lepton plus Jet Channel Using $\int \mathcal{L}dt = 9.4$ fb^{-1} CDF Data

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: August 9, 2013)

We present a measurement of the cross section of single top s-channel process. For this process, we are looking for events consistent with the signature of a charged lepton (electron or muon), large missing transverse energy ($\not\!\!E_T$), and two jets, of which at least one is required to originate from a bottom quark. This measurement uses the full CDF [1] dataset, corresponding to an integrated luminosity of 9.4 fb^{-1}. Discrimination between the single top s-channel signal and the comparatively large backgrounds is improved through the use of a multivariate technique in the form of a TMVA [2] trained Neural Network (BNN) discriminant.

We measured the single top s-channel cross section to be $\sigma_{s-\mathrm{channel}} = 1.41^{+0.44}_{-0.42}$ pb, assuming the top quark mass is 172.5 GeV/c^2. This is compatible with standard model prediction. The result corresponds to a significance of 3.8σ over the background-only hypothesis.

I. INTRODUCTION

In Standard Model, top quark could either be produced in pair or singly. In 1995, top quark was first discovered in top quark pair production which is determined by strong interaction. The single top quark production process determined by electro-weak interaction was discovered later on, simultaneously by CDF and D0 experiment in 2009 [3][4]. At Tevatron, the t-channel and s-channel process are dominant in single top production. The predicted cross section for t-channel is 2.10 ± 0.19 pb, and the predicted cross section for s-channel is 1.05 ± 0.07 pb, the cross section for Wt-channel is negligible [5].

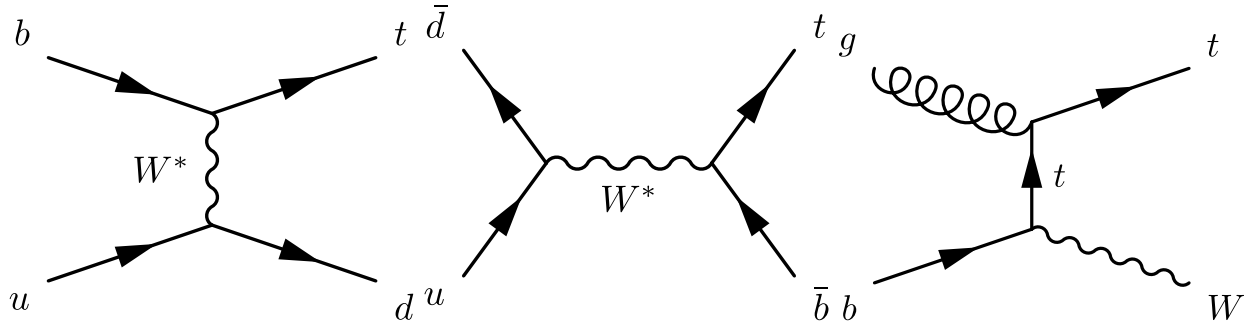

FIG. 1: Feynman diagram for the single top process. The left one is t-channel process, the center one is s-channel process and the right one is Wt-channel process.

Within standard model, the single top process provides the only opportunity to directly measure the V_{tb} element in Cabbibo-Kobayashi-Maskawa (CKM) matrix. Furthermore, since the top quark is the only quark that decays before hadronization, this allows us to study the polarization property of the top quark. The single top process is also sensitive to several beyond standard model processes. Moreover, since Large Hadron Collider(LHC) is pp collision, this s-channel process would be very difficult for them to analysis. This Tevatron results would be the best results for many years.

In this note, we described our measurement of single top s-channel process cross section corresponding to CDF full dataset. Our selection of event samples are described in Section II. To further improve the sensitivity of our analysis, we used TMVA to train a neural network to separate single top s-channel events from other backgrounds. This is described in Section IV. This analysis is sharing the same event selection strategy and background model method as the WH analysis [6] at CDF.

II. DATA SAMPLE AND EVENT SELECTION

In this analysis, we select events consistent with a W-boson decay plus two energetic b-quark jets. The W-boson events are selected by requiring a single, isolated lepton with $p_T > 20$ GeV/c, and the presence of $\not\!\!E_T$ in the event with a different threshold depending on the quality of the lepton reconstruction algorithm in use.

We use four different lepton reconstruction algorithm, and we divide them into following categories to keep them orthogonal to each other.

- CEM: central tight electron
- CMUP and CMX: central tight muon
- Extented Muon Category(EMC): loose muons and reconstructed isolated track lepton candidates

The data used for this analysis are collected from different datasets. CEM events are from high-p_T electron dataset, CMUP and CMX events are from high-p_T muon dataset, while EMC events are collected using a set of triggers based on the missing transverse energy($\not\!\!E_T$) and jet information.

To look for the W boson in the events, we require a charged lepton p_T to be larger than 20 GeV/c and $\not\!\!E_T > 10$ GeV for CMUP and CMX events, $\not\!\!E_T > 20$ GeV for CEM and EMC events.

Jets information used in the analysis are reconstructed with the JetClu algorithm with a cone size of 0.4. The jet energy is corrected for the η-dependence of the calorimeter, for multiple $p\bar{p}$ interactions and absolute energy scale. Tight jets are required to have corrected $E_T > 20$ GeV/c^2 and $\eta < 2.0$. Only events with exactly two tight jets are accepted.

We also employ b-tagging algorithms to further select our events. A b-tagging algorithm denoted as The HOBIT [7] is used in the analysis. We require at least one of the jets to be tagged by HOBIT. We defined two operational points of the tagging algorithm based on the output value of HOBIT, Tight, output larger than 0.98 and Loose, output larger than 0.72.

Based on the tagging information, we divide events into following four orthogonal tagging categories:

- TT: Exactly two jets tagged by HOBIT Tight
- TL: One jet is HOBIT Tight, another jet is HOBIT Loose, but not HOBIT Tight
- T: One of the jets is HOBIT Tight, other jets are not HOBIT Loose
- LL: None of the jets are HOBIT Tight, exactly two jets are HOBIT Loose

Multijet QCD background is further suppressed with the use of a multivariate technique developed by Support Vector Machine [8].

The Monte Carlo sample used in this analysis to model the background and signal events are generated by different Monte Carlo packages. Single top samples are generated by POWHEG [9] and showered by PYTHIA [10]. For W + jets and Z + jets samples are modeled by ALPGEN [11]and showered with PYTHIA. For $t\bar{t}$ and diboson samples are generated with PYTHIA.

III. BACKGROUNDS ESTIMATION

We determine the fraction of W + jets events for each lepton category by fitting the $\not\!\!E_T$ distribution of pretag samples. For single top, $t\bar{t}$, diboson and Z + jets samples are normalized to the theoretical expectation, while W + jets and multijet QCD samples normalization are free to float in the likelihood fit.

The b-tagged W + heavy flavor sample is extracted from the total W + jets pretag sample by applying the heavy flavor fraction, which calculated by using inclusive W + jets sample, and b-tagging efficiency.

The W + Mistag sample in the tagged region are calculated from pretag W + light flavor samples by applying a per-jet false tag rate parameterization(Mistag matrix).

The normalization of multijet QCD background in tagged region are calculated from the multijet background in pretag region and apply the tagging efficiency calculated from number of data events in each tagging category divided by number of events in pretag region. The shape of multijet background in tagged region are determined by the shape in pretag region.

The prediction for number of events in each tagging category are shown in Table I.

Category	TT	TL	T	LL
WW	1.7±0.4	13.2±2.7	184±23	24.8±3.9
WZ	17.8±2.2	21.2±2.0	52.7±5.4	9.9±0.9
ZZ	2.4±0.3	2.4±0.2	7.1±0.7	0.96±0.08
Z + jets	10.9±1.2	20.7±2.3	163±18	27.1±3.1
$t\bar{t}$	163±21	194±19	502±50	58.1±6.6
Higgs	6.1±0.6	6.4±0.4	10.3±0.7	1.7±0.2
Wbb	246±99	327±130	1166±468	109±44
Wcc	19.0±7.8	120±49	1158±467	164±67
W + Mistag	4.3±1.3	62±13	978±141	242±34
Multijet	29±12	47±19	281±112	45±18
t and Wt-channel	18.1±2.5	35.3±4.2	251±28	13.6±1.5
s-channel	54.5±6.7	61.2±5.6	109±10	17.8±2.1
Total Prediction	573±155	911±248	4860±1320	714±181
Observed	466	765	4620	718

TABLE I: Summary of background and signal prediction in each tagging category, with systematic uncertainties of cross section included.

IV. FINAL DISCRIMINANT

To further separate the signal from background, and increase the sensitivity of this analysis, we use TMVA package trained a neural network to be the final discriminant.

The input variable used in this neural network training are summarized in the Table II. We trained separate neural networks for each tagging category. The Validation plot of the modeling of input variable and output value for each neural network in pretag region are shown in appendix.

variable	Tight Lepton	EMC
$M_{l\nu b}$	√	√
$M_{l\nu bb}$	√	√
Lep p_T	√	√
M_{jj}	√	√
$\cos\theta_{lj}$	√	√
H_t	√	√
$M^T_{l\nu b}$		√
b jet selector output	√	

TABLE II: Summary of input variables used for final discriminant for different lepton category

$M_{l\nu b}$ The reconstructed top quark mass

$M_{l\nu bb}$ The reconstructed mass of the charged lepton, $\not\!\!E_T$ and two jets

Lep p_T The p_T of the charged lepton

M_{jj} The reconstructed mass of two jets corrected using neural network [12]

$\cos\theta_{lj}$ The cosine of the angle between the charged lepton and the jet selected to reconstruct top quark in the top quark rest frame

H_t The scalar sum of transverse energy of the charged lepton, $\not\!\!E_T$ and all jets

$M^T_{l\nu b}$ The transverse mass of the reconstructed top quark

b **jet selector output** The output of the neural network used to select the b jet originated from top quark

The final discriminant output distributions of each tagging category are shown in Figure 2 and Figure 3.

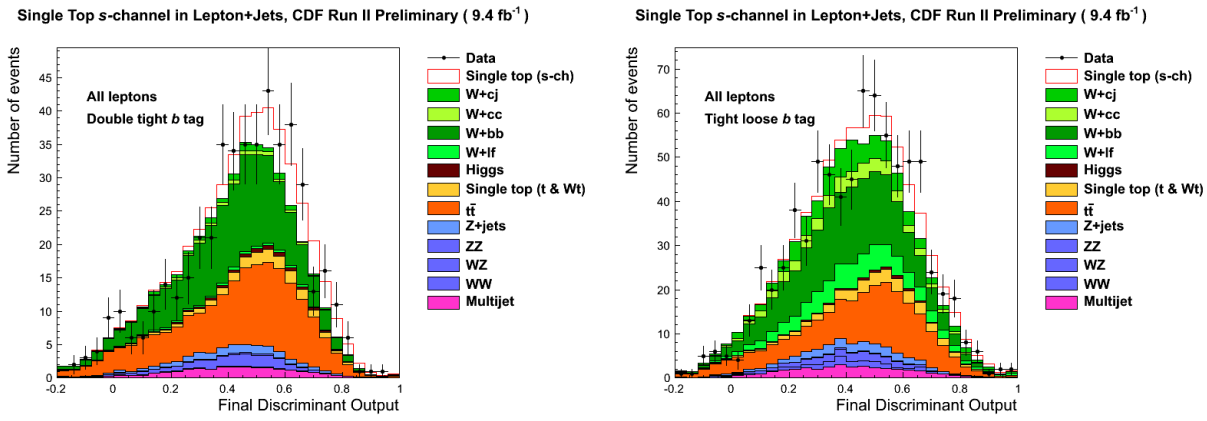

FIG. 2: Final discriminant output for each tagging category and all lepton category combined

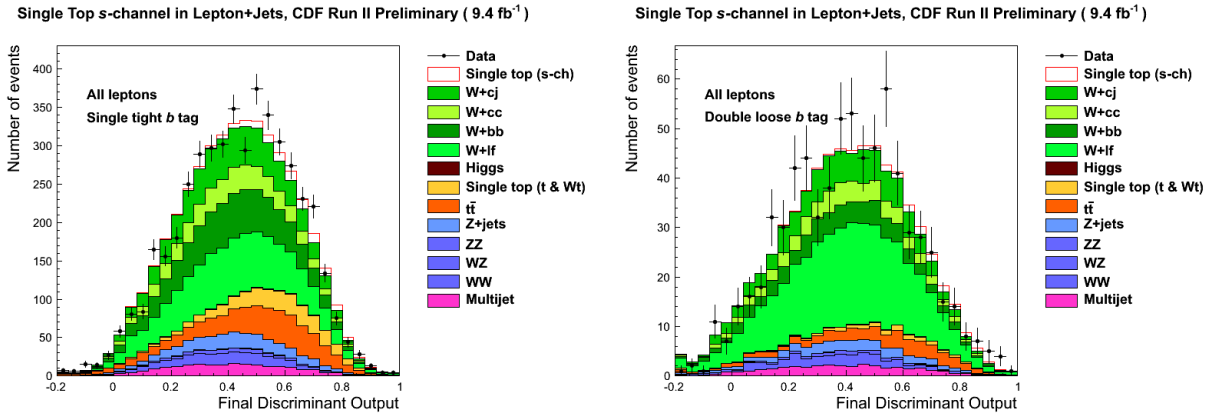

FIG. 3: Final discriminant output for each tagging category and all lepton category combined

Source of uncertainty	Rate	Shape	Affected samples
b tagging scale factor uncertainty	4%-18%		$t\bar{t}$, single top, WZ, ZZ, Higgs
Charm mistag rate	7%-37%		WW
W+jets mistag rate	4%-37%		W + Mistag jets
Luminosity uncertainty	6%		$t\bar{t}$, single top, diboson, Higgs
Lepton acceptance uncertainty	2%-4%		$t\bar{t}$, single top, diboson, Higgs
Cross section uncertainty	6%-10%		$t\bar{t}$, single top, diboson, Higgs
Initial/Final state radiation	0%-10%	$\surd$	$t\bar{t}$, single top
Multijet normalization	40%		Multijet
Z+jets normalization	45%		Z+jets
Wbb and Wcc normalization	30%		Wbb, Wcc
Wc normalization	30%		Wc
Jet energy scale	0%-10%	$\surd$	All
Normalization and factorization scale		$\surd$	W+jets
Electron multijet background		$\surd$	Electron multijet

TABLE III: Summary of all systematics considered in this analysis

V. SYSTEMATIC UNCERTAINTIES

The summary of all the systemactic uncertainty considered in this analysis are listed in the Table III.

VI. MEASUREMENT

We measure the single top cross section using a Bayesian binned likelihood technique [13] assuming a flat prior in the cross section and integrating the posterior over all sources of systematic uncertainty.

The posterior probability distribution of single top s-channel cross section from calculation is shown in Figure 4. From the distribution, we measure the s-channel cross section to be $\sigma_{s-\mathrm{channel}} = 1.41^{+0.44}_{-0.42}$ pb.

We also measured the p-value by generating pseudo-experiment with both background only and signal plus background hypothesis, as shown in Figure 5. The p-value for the observed cross section is 0.0000597, which corresponds to a significance of 3.8σ.

As a sanity check, we also measured the single top s-channel cross section in each lepton and tagging category, it shows that the results are compatible with each other between categories. As shown in Figure 6.

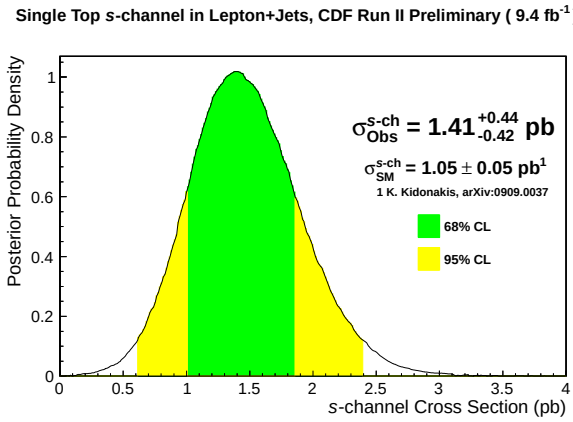

FIG. 4: Posterior probability density distribution for the cross section measurement. The green and yellow regions represent the smallest intervals enclosing 68% and 95% of the posterior density integrals.

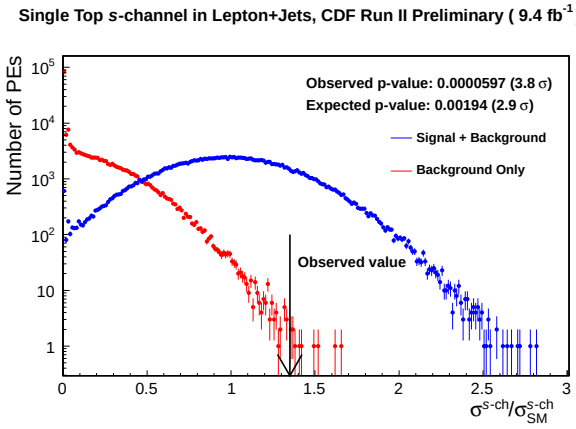

FIG. 5: Pseudo-experiment output with background only hypothesis and background plus signal hypothesis. The arrow is pointing at the measured cross section value.

VII. CONCLUSION

We have presented the results of a search for the single top s-channel. We find that for the dataset corresponding to integrated luminosity of 9.4 fb^{-1}, the data agrees with the Standard Model background predictions within the systematic uncertainties.

We measure the single top s-channel cross section to be $\sigma_{s-\text{channel}} = 1.41^{+0.44}_{-0.42}$ pb, assuming the top quark mass is 172.5 GeV/c^2. This is compatible with standard model prediction and is also compatible with previous CDF measurement. This corresponds to a significance of 3.8σ.

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the

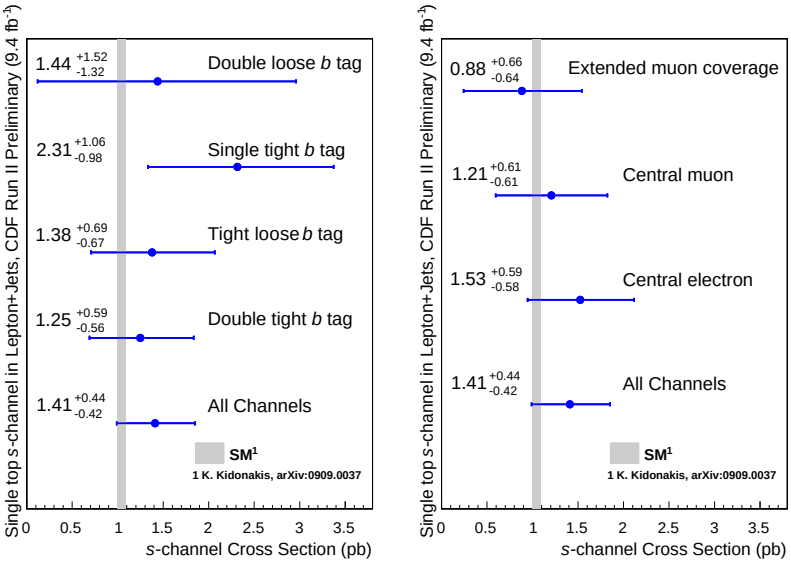

FIG. 6: Single top s-channel cross section measured in different tagging and lepton channels. The grey line on the plot is the theoretical cross section value with uncertainties.

Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; and the Australian Research Council (ARC).

[1] F. Abe *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988);
D. Amidei *et al.*, Nucl. Instum. Methods Phys. Res. A **350**, 73 (1994);
F. Abe *et al.*, Phys. Rev. D **52**, 4784 (1995);
P. Azzi *et al.*, Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995);
The CDF II Detector Technical Design Report, Fermilab-Pub-96/390-E.
[2] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," PoS A CAT 040 (2007) [physics/0703039].
[3] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, 092002 (2009).
[4] V.M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **103**, 092001 (2009).
[5] N. Kidonakis, arXiv:0909.0037 [hep-ph] (2009).
[6] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **109**, 111804 (2012).
[7] T. Aaltonen *et al.* (CDF Collaboration), CDF Public Note 10803 (2011).
[8] F. Sforza, V. Lippi, G. Chiarelli, J. Phys.: Conf. Ser. **331** 032045 (2011).
[9] P. Nason, JHEP **0411** 040, (2004) S. Frixione, P. Nason and C. Oleari, JHEP **0711** 070, (2007). S. Alioli, P. Nason, C. Oleari and E. Re, JHEP **1006** 043, (2010). S. Alioli, P. Nason, C. Oleari and E. Re, JHEP **0909** 111, (2009). E. Re, Eur. Phys. J. **C71** 1547, (2011).
[10] P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, E. Norrbin, and T. Sjöstrand, Computer Phys. Commun. **135** (2001) 238.
[11] M.L. Mangano *et al.*, JHEP 0307:001, 2003.
[12] T. Aaltonen *et al.*, Improved b-jet Energy Correction for $H \to b\bar{b}$ Searches at CDF, arXiv:1107.3026 (2011).
[13] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D86, 010001 (2012).

APPENDIX A: FINAL DISCRIMINANT INPUT VARIABLES

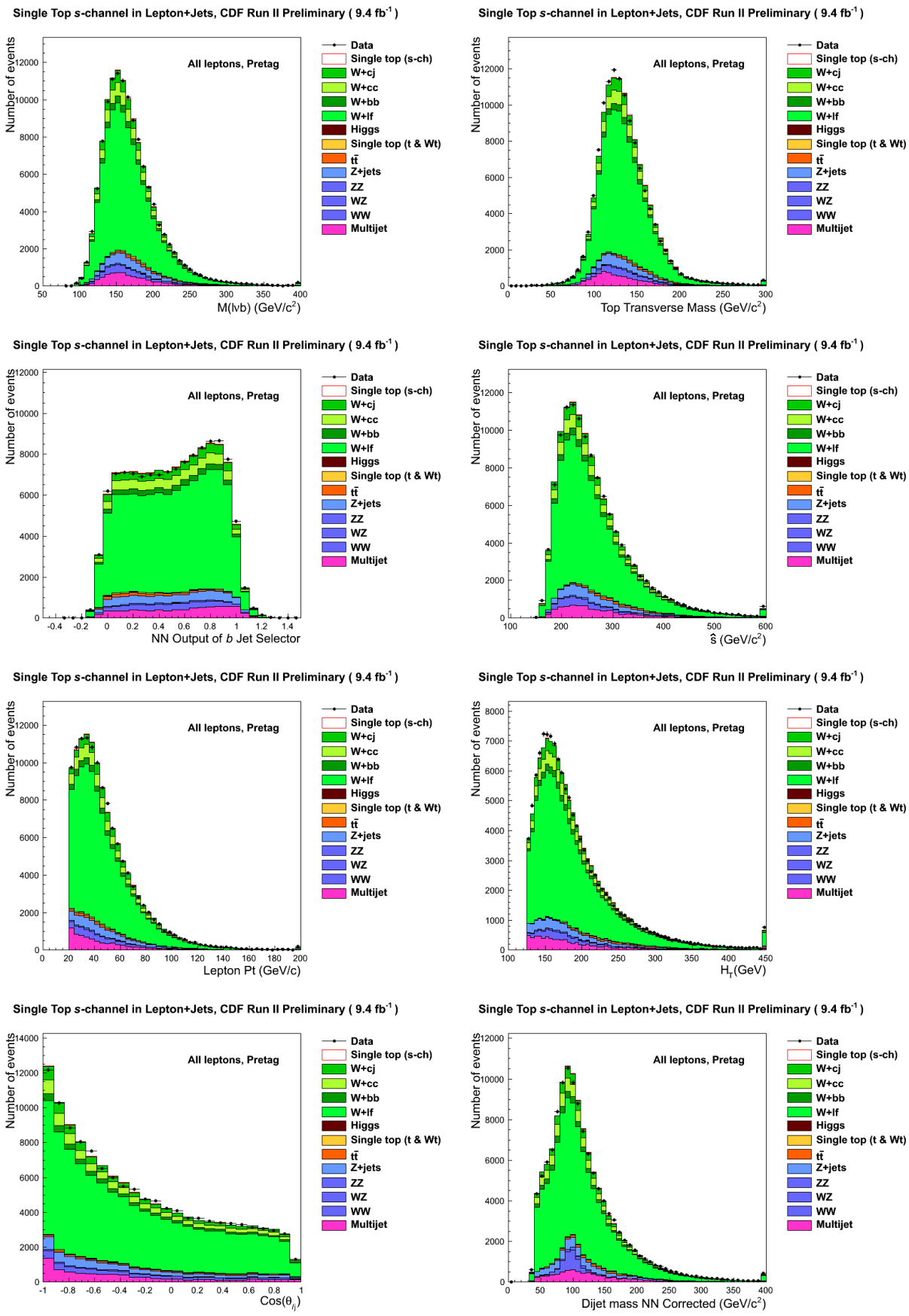

FIG. 7: Neural network input variables in pretag region

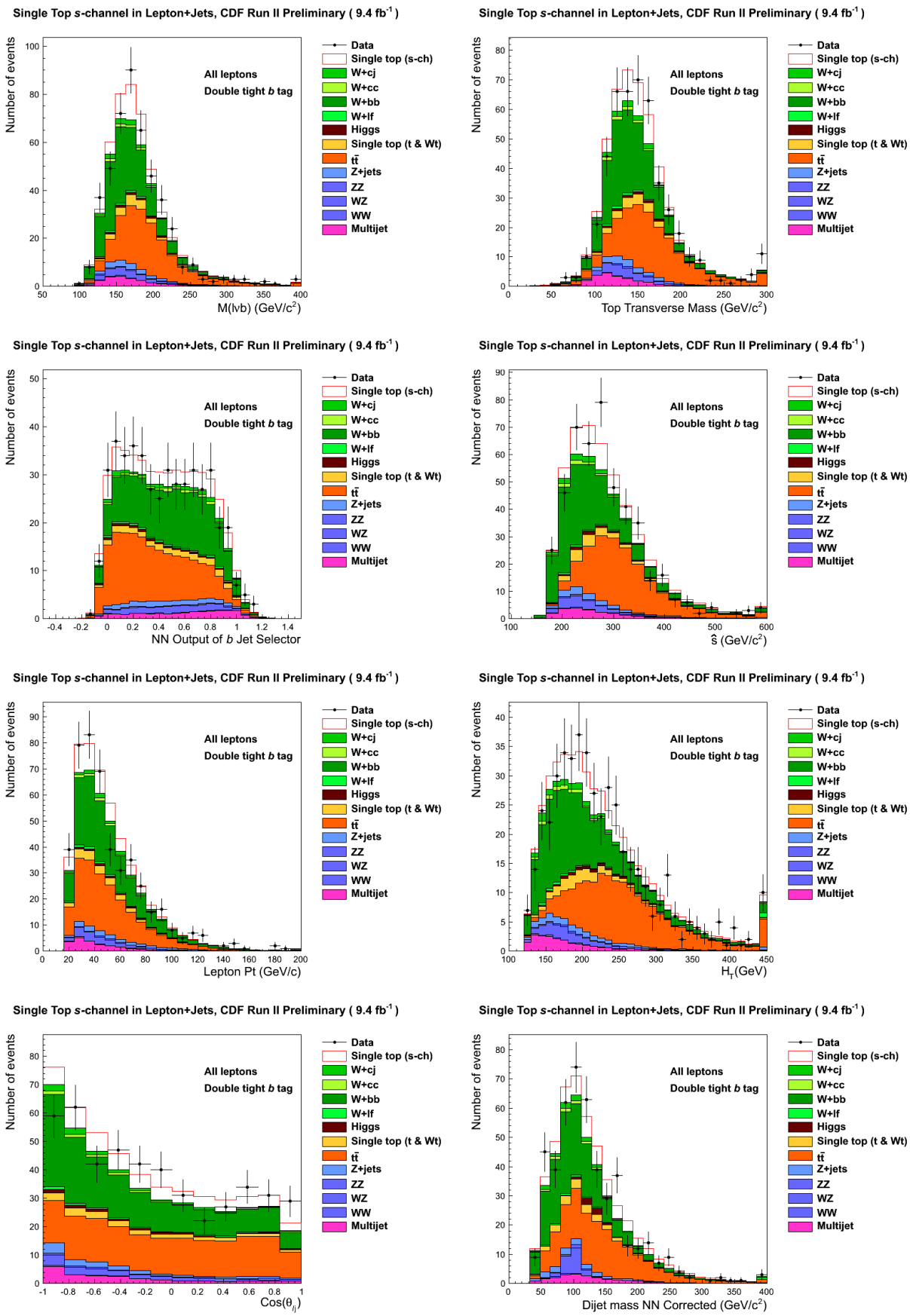

FIG. 8: Neural network input variables in TT region

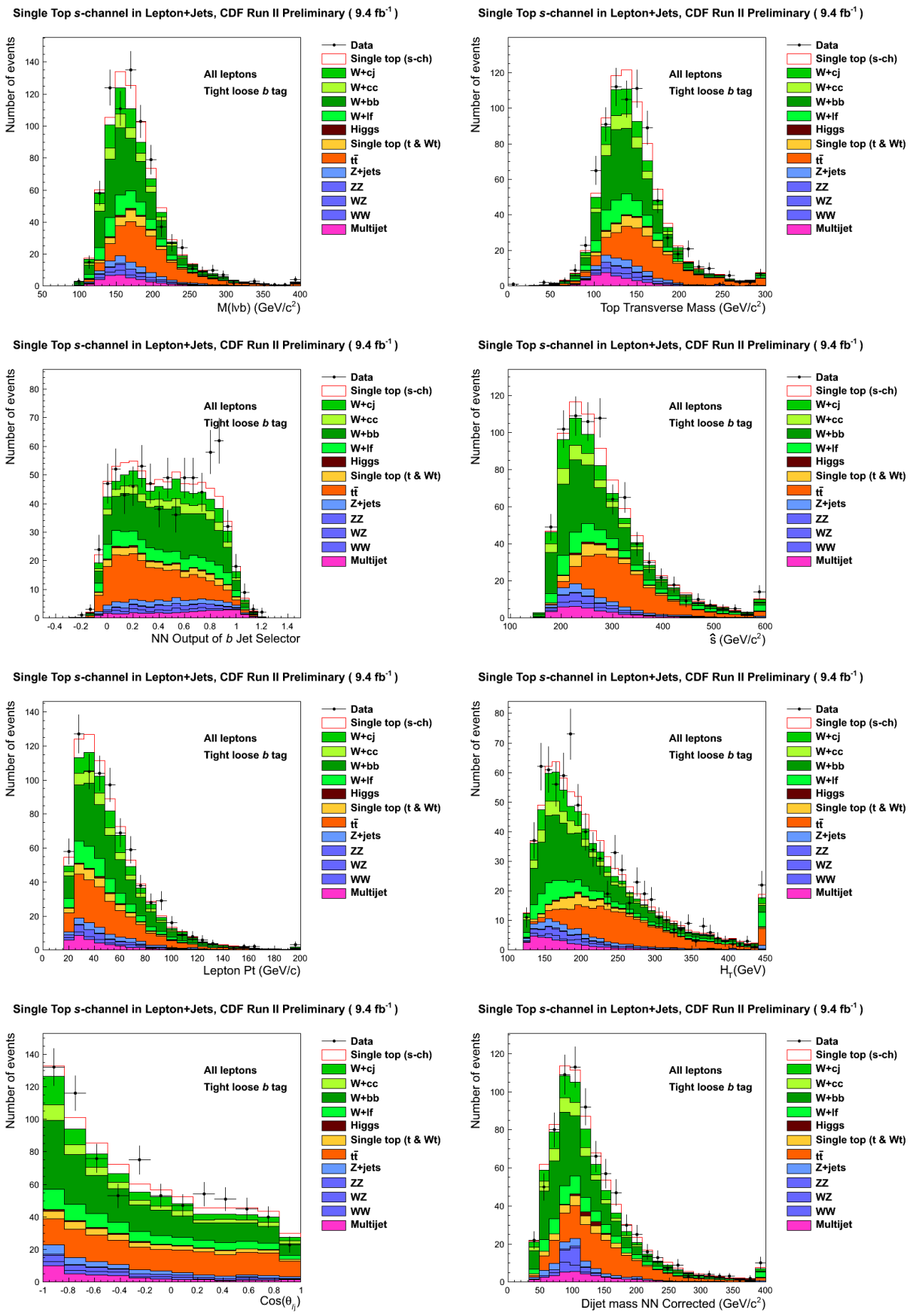

FIG. 9: Neural network input variables in TL region

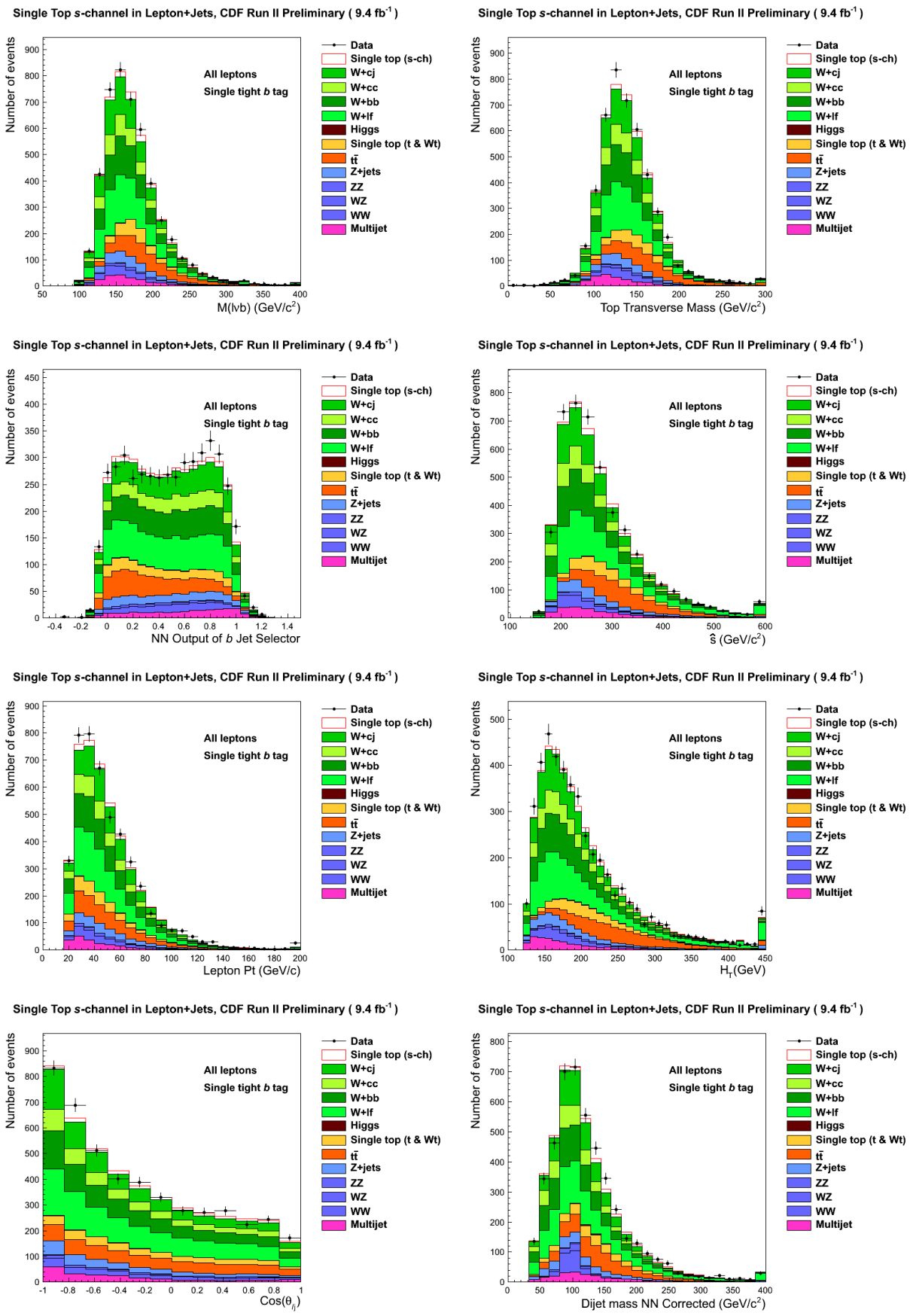

FIG. 10: Neural network input variables in T region

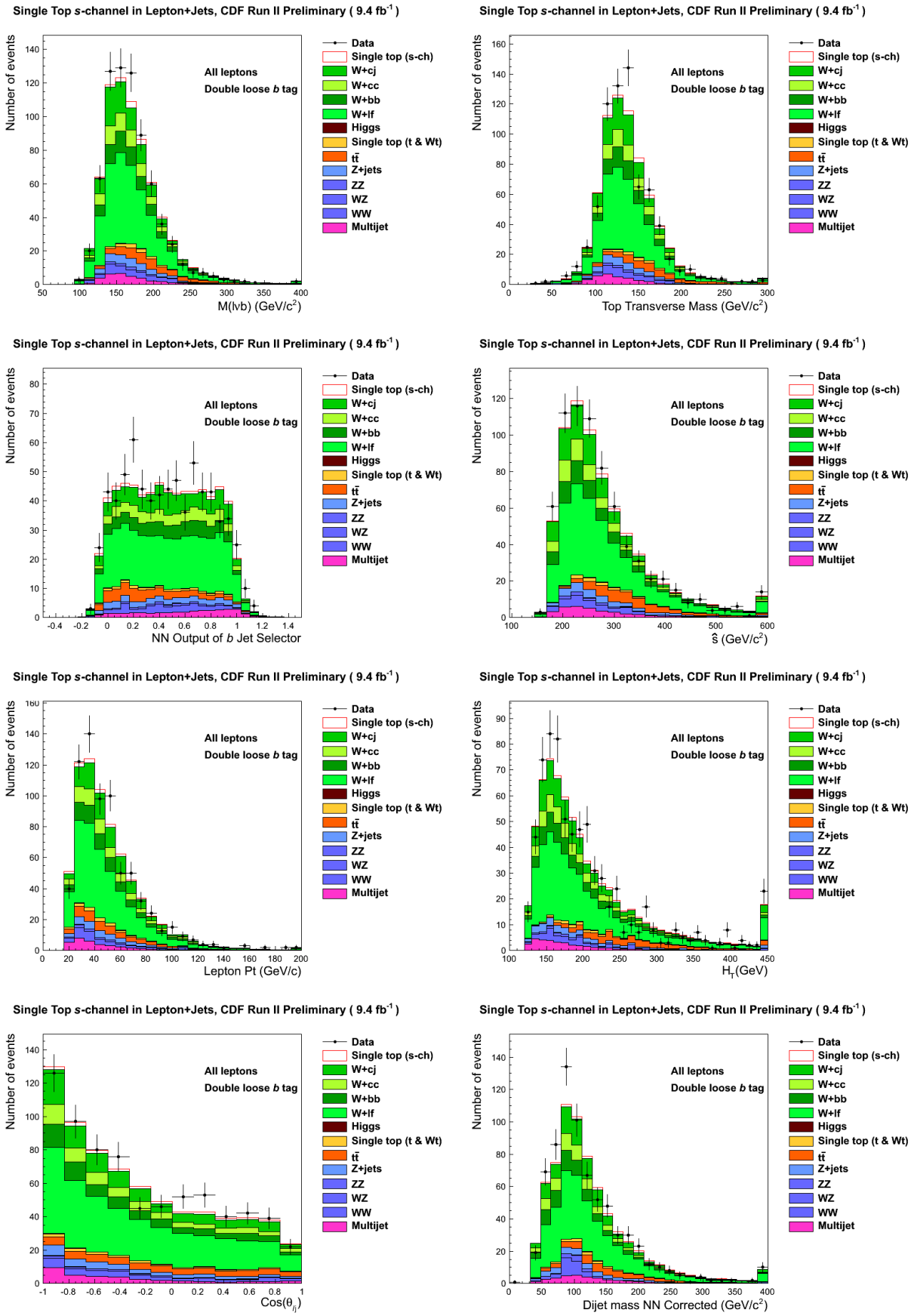

FIG. 11: Neural network input variables in LL region

APPENDIX B: FINAL DISCRIMINANT OUTPUT

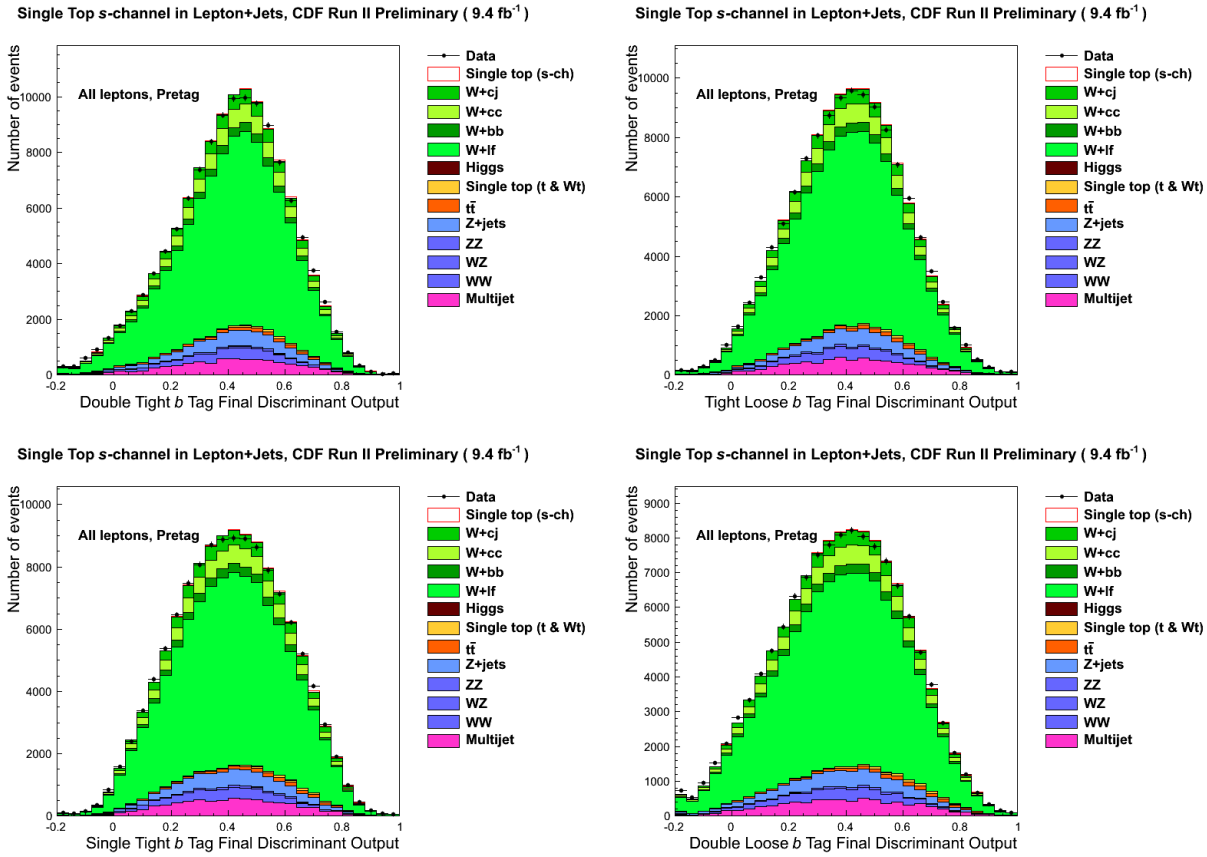

FIG. 12: Neural Network Output in Pretag region